

# The effects of performance expectation and question difficulty on text study time, response certitude, and correct responding

WILLIAM A. STOCK, KRISTEN S. WINSTON,  
JOHN T. BEHRENS, and MARIA HARPER-MARINICK  
*Arizona State University, Tempe, Arizona*

On the basis of a premise that persons modify studying behavior in order to maintain response certitude at acceptable levels, performance expectations and question difficulty were varied factorially in two studies, and their effects on study time, correct responding, and response certitude were evaluated. The manipulations did not affect study time, but correct responding was in the anticipated direction in both studies, significantly so in Experiment 2. For response certitude, an expected interaction of performance expectation  $\times$  question difficulty  $\times$  trial was in the anticipated direction in both studies, significantly so in Experiment 1.

In studies of the effects of feedback on learning and retention, Kulhavy, Yekovich, and Dyer (1976, 1979) noted a consistent, direct relation between response certitude (hereafter *certitude*) and the acquisition of knowledge when certitude was defined as a subjective estimate of the likelihood of responding correctly. Also, feedback study time was related directly to certitude for persons making an incorrect response, and inversely related for those making a correct response.

We believe certitude is a control variable operating at a high level of cognitive organization. In this view, study behavior is controlled by the discrepancy between a reference level of certitude and an estimated level associated with current studying. Operating as a continuous, ongoing, negative feedback loop, the object of this control mechanism is to reduce and maintain the discrepancy within acceptable bounds. We propose that individuals will, given the opportunity, continue studying when estimated certitude is less than its reference value, and, further, will cease studying when estimated certitude equals or exceeds the reference value.

We assume that the reference value is determined by both individual and environmental factors. Thus, interests and prior knowledge, general levels of aspiration, and history of studying behavior will all influence a target level for certitude. Some situational factors may also play a role in determining this reference value. Externally suggested performance standards represent one situational factor with implications for instruction. Estimated certitude is assumed to be a function of situational factors related to the study task, such as text and question difficulty.

In two studies, we directly manipulated performance expectation—an attempt to elevate the reference level for

certitude prior to studying. In both studies, some subjects were asked to study a text until they felt they could correctly answer all subsequent questions about the text, and other subjects were asked to study until they felt they could correctly answer at a minimal level of proficiency. Experiments 1 and 2 differed in the type of opportunity for continued study. In Experiment 1, the subjects had one opportunity to study the text as long as they wished; in Experiment 2, the text was presented for a fixed time, and the subjects chose to reread it as many times as they wished.

Our expectation was that study times in the high performance expectation conditions would exceed those in the minimal performance expectation conditions. This expectation was based upon the belief that all subjects would incorporate the externally suggested performance standard in determining a reference value for certitude, thereby creating a higher standard for the subjects in the high expectation conditions.

We also systematically varied the difficulty level of the questions asked after the subjects had completed study. If studying is restricted to a period prior to the presentation of questions, then differences in question difficulty should lead to orderly changes in certitude. Here, after the subjects had studied for as long as they wished, the text was removed and the subjects answered either four harder or four easier questions. Although the instructions indicated a standard of answering performance, the subjects were not informed about question difficulty.

Concerning judgments of certitude, we expected an interaction between performance expectation, question difficulty, and trial. Because the questions themselves provided information regarding the likelihood of correct responding, we anticipated that the subjects who received minimal performance expectations and the more difficult questions would be most likely to adjust estimates of their certitude as they encountered the more difficult questions.

Correspondence may be addressed to William A. Stock, Division of Psychology in Education, Arizona State University, Tempe, AZ 85287-0611.

Relative to other subjects, these subjects should have exhibited a flatter (or more negative) certitude function across the question trials.

## EXPERIMENT 1

### Method

**Subjects.** Sixty-four undergraduate volunteers from introductory educational psychology classes at Arizona State University were randomly assigned in equal numbers to one of four treatment combinations. They received class credit for participation.

**Design.** Performance expectation and question difficulty were between-subjects factors, and question trial (four levels) was a within-subjects factor. Performance expectation and question difficulty were set at high and low levels.

**Materials.** Materials included instructions, one practice and one target text passage, a set of four hard and four easy questions for the target passage, an item soliciting subjects' certitude judgments, and a set of four math problems used for interpolated activity between text processing and responding. The practice passage was a 288-word segment on solar energy, and the target passage was a 179-word segment on wilderness medicine, both of which were written at a 12th-grade reading level. The solar energy passage was used to obtain an estimate of reading time.

The math problems (a column addition of four two- and three-digit numbers, a subtraction involving a six- and a five-digit number, a multiplication of a three-digit number by a two-digit number, and a long division of a five-digit number by a three-digit number) were used to reduce short-term memory effects and were presented to the subjects after exposure to the target passage and prior to questioning.

Four five-alternative, multiple choice questions covering the wilderness medicine passage were our hard questions, and corresponding, simpler, four-alternative multiple choice questions were constructed as our easy questions. The target passage and difficult questions were constructed and normed in an unpublished thesis by the second author. The easier items were constructed by simplifying the response options, making the correct option somewhat more salient, and deleting one incorrect option. The certitude item, presented prior to each question and after the last question (five occasions), had 10 alternatives varying from 1 (no confidence) to 10 (complete confidence).

**Procedure.** An IBM BASIC program controlled presentation of all instructions and materials, except the math problems, and recorded all responses and text study times for the two passages. Four IBM AT computers with color monitors were set up in a quiet well-lighted room, one station in each corner. Each computer was programmed for one of the four between-subjects treatment conditions. The subjects were randomly assigned to one of the four computers in order of their appearance, with a maximum of four participants every 15 min. Prior to initiating the program, the subjects were briefed on the use of the keyboard. Prior to and during the use of the program, all questions were individually answered.

Order of presentation was fixed. An introduction described the study, followed by instructions to read the solar passage as quickly as possible with complete comprehension. After reading the solar energy passage, the subjects received instructions to read the target passage. These instructions contained the performance expectation and indicated that studying was subject-paced. Under high performance expectation, a statement suggested studying so as to ensure correct responding to all questions that might be asked about the text. Under low performance expectation, a statement suggested studying so as to ensure correct responses on at least 25% of the questions. After studying the text, each subject was requested to turn over a sheet of plain white paper next to the computer and work the mathematics problems. Instructions at the bottom of the paper directed the subjects to press the space bar on the keyboard to continue. Finally, five certitude items and four questions were presented alternately.

### Results and Discussion

Passage reading times were transformed using logarithms (Winer, 1971), and preliminary analysis indicated

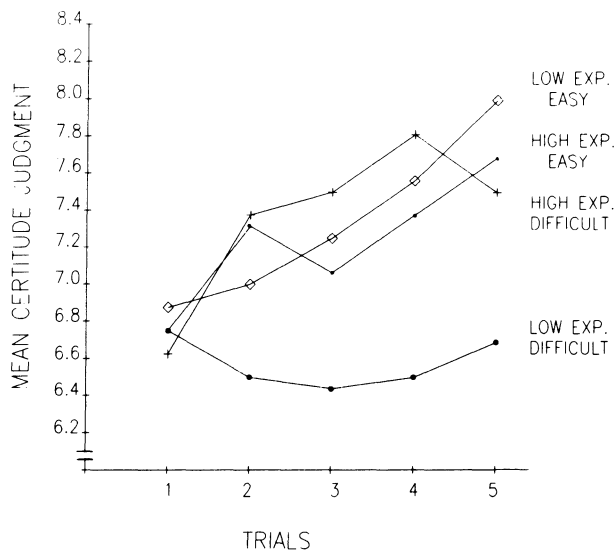


Figure 1. Mean certitude judgment as a function of performance expectation, question difficulty, and trial.

comparable reading rates among the four groups. Other analyses indicated that using reading time for the solar passage as a covariate was unnecessary. Consequently, reading time for the solar energy passage is not considered further.

The sum of correct responses was analyzed by a  $2 \times 2$  (expectation  $\times$  difficulty) analysis of variance. Neither performance expectation nor the interaction of performance expectation and question difficulty was significant. Although not significant, correct responding was higher in the high performance expectation conditions. Question difficulty had a significant effect on performance [ $F(1,60) = 10.56, p < .01$ ], with hard and easy forms of questions having means of 2.78 and 3.58 items correct, respectively.

Completed next was a  $2 \times 2$  analysis of variance involving log reading time of the target passage. There were no significant effects. The absence of an effect for performance expectation was contrary to our hypothesis.

Certitude judgments were analyzed, using a  $2 \times 2 \times 5$  analysis of variance design, with the five occasions of measurement forming a repeated measures factor. Paralleling the results of the analysis of study time, effects of performance expectation, question difficulty, and their interaction were not significant.

A trend analysis of the occasion factor indicated that the linear component of trials was significant [ $F(1,60) = 19.76, p < .01$ ], as was its interaction with performance expectation and question difficulty [ $F(1,60) = 4.33, p < .05$ ]. Examination of Figure 1 will suggest a general linear increase in mean confidence levels across occasions for subjects in three conditions, and, as anticipated, a flat line in the low performance expectation/hard questions condition.

A quadratic component of the performance expectation  $\times$  trial interaction also was significant [ $F(1,58) = 8.77, p < .01$ ], and reexamination of Figure 1 will suggest that a combined curve for high performance expectation con-

ditions would be concave downward, whereas a combined curve for low performance expectation conditions would be concave upward. The subjects in the high performance expectation conditions reached a peak of response confidence, with respect to answering this set of questions, earlier than did the subjects in the low performance expectation conditions.

## EXPERIMENT 2

Experiment 2 differed from Experiment 1 in two ways. First, instructions for studying the text segment were modified. The first occurrence of the performance expectation instructions was typed in bold capital letters, and followed by a repetition of the performance expectations several lines below the body of the instructions. This alteration was intended to increase the salience of the performance expectation. A second difference involved presenting to-be-studied text for a fixed time period. This period was set at 55 sec, a value 1 sec above the minimum reading time observed for this passage in Experiment 1. After expiration of a presentation period, the subjects were given an opportunity to re-view the passage. This alteration introduced a choice point at which subjects with higher performance expectations might differentiate themselves from those with lower performance expectations.

### Method

**Subjects.** Seventy-seven undergraduate volunteers from introductory educational psychology classes at Arizona State University were randomly assigned to high expectation/difficult question ( $n = 19$ ), high expectation/easy question ( $n = 18$ ), low expectation/difficult question ( $n = 21$ ), and low expectation/easy question ( $n = 19$ ) conditions. The subjects received class credit for participation.

**Design and Materials.** The design and materials of Experiment 2 replicated those of Experiment 1.

**Procedure.** The procedure varied from that of Experiment 1 as noted above. After presentation of the text study instructions, the target passage was displayed for 55 sec, and immediately followed by a screen that asked if the subject wished to re-view the passage. A "yes" or "no" response was required. As in Experiment 1, at any time during the presentation of the passage, the subjects could discontinue presentation by pressing the space bar.

### Results and Discussion

In a  $2 \times 2$  analysis of variance using total corrects, performance expectation [ $F(1,73) = 4.36, p < .05$ ] had a significant effect, while both question difficulty and the interaction of performance expectation and question difficulty were not significant. The mean numbers of items correct for high and low performance expectation conditions were 3.30 and 2.83, respectively—a result consistent with the initial hypothesis regarding manipulation of expectation. Although question difficulty was not statistically significant, hard and easy forms of questions had means of 2.89 and 3.24 items correct, respectively.

Completed next were separate  $2 \times 2$  analyses of variance involving log reading time of the wilderness passage and number of re-views as dependent variables. The ef-

fects of expectation, question difficulty, and their interaction were not significantly related to either variable. As in Experiment 1, the absence of an effect for performance expectation is contrary to our hypothesis. Although performance expectation manifests itself in correct responding, it does not affect time spent studying.

Completed next was a  $2 \times 2 \times 5$  mixed analysis of variance with certitude judgment as the dependent variable. Trials were partitioned into a one degree of freedom component testing a linear trend, and a three degrees of freedom component pooling the quadratic, cubic, and quartic effects. Effects were generally not significant, except for the linear component of the performance expectation  $\times$  question difficulty  $\times$  trial interaction, which approached statistical significance [ $F(1,73) = 2.31, p < .08$ ].

Because this effect had theoretical significance, we examined the plot of the three-factor interaction. Curves for three conditions (high expectation/hard questions, high expectation/easy questions, and low expectation/easy questions) all tended to rise, with the high expectation/easy question condition consistently higher than all other conditions, whereas means in the low expectation/difficult question condition first rose and then fell below all other conditions.

## GENERAL DISCUSSION

Informal postexperiment discussions with some of the subjects suggested that the low performance expectation conditions may have been ineffectual—some subjects in these conditions reported studying the text as they would any other academic material. At this point, we would argue for attempting other manipulations directed at altering the reference level for certitude. Such manipulations would have to be constructed in a fashion that avoids merely instructing the subject to study longer. Our choice is to evaluate our underlying rationale in other experimental settings—a choice based on other, more encouraging results.

That correct responding differed in the anticipated direction as a function of performance expectation (significantly so in Experiment 2) was consistent with our hypothesis. Because these studies were completed within very short time frames (no more than 20 min for any subject) with a minimal variation in experimental instruction, we believe that more intensive efforts to induce differential performance expectations can have a more robust effect. Similarly, the use of more extensive materials would provide a more realistic setting for examining our position.

Regarding judgments of certitude, the results of Experiment 1 conformed exactly with our prior expectations, and the results of Experiment 2 were consistent with our expectation, although they did not reach conventional significance levels. Thus, there is evidence that judgments of response certitude change in the predicted manner.

## REFERENCES

- KULHAVY, R. W., YEKOVICH, F. R., & DYER, J. W. (1976). Feedback and response confidence. *Journal of Educational Psychology*, *68*, 522-528.
- KULHAVY, R. W., YEKOVICH, F. R., & DYER, J. W. (1979). Feedback and content review in programmed instruction. *Contemporary Educational Psychology*, *4*, 91-98.
- WINER, B. J. (1971). *Statistical principles in experimental design*. New York: McGraw-Hill.